DEVELOPMENTAL TESTBED CENTER (DTC)

Zoltan Toth[1]

Jamie Wolf[2], Ligia Bernardet[3], Hui Shao[2], Isidora Jankov[4], Tressa Flower[2]

Developmental Testbed Center

[1] Global Systems Division (GSD) NOAA/OAR/ESRL
[2] NCAR
[3] CIRES at GSD
[4] CIRA at GSD

Acknowldegements:
Bill Kuo, Louisa Nance, Barbara Brown, Scott Hausman, and Steve Koch

4th NOAA Testbed Workshop, April 2-4, 2013

OUTLINE / SUMMARY

- **Overview**
  - Transition of research into operations
    - For Numerical Weather Prediction (NWP)

- **Research to Operations (R2O) testing**
  - WRF, HWRF, GSI, SREF, supported by
    - Operations to Research (O2R, e.g., code repositories)

- **Outlook**
  - Discussions on scope of DTC
    - Improve current & next generation NWP systems
    - New Cooperative Agreement
  - Build modern NWP IT Environment (NITE)
  - Strengthen collaboration with other NOAA testbeds & programs

BACKGROUND

- **History**
  - Initiated in 2004; NOAA funding increases in 2009 & 2010
- **Organization**
  - Interagency level – Charter – Bill Kuo, Director
    - NOAA, NSF, NCAR, USAF
  - NOAA level
    - OAR-GSD, HFIP, USWRP, with EMC support
- **Staffing**
  - NCAR/RAL – Under NOAA Cooperative Agreement
  - ESRL/GSD
- **NOAA Cooperative Agreement**
  - Present - NCAR, 2008-2013
  - Next phase – 2014-2019
    - Announcement of Opportunity being prepared
      - Competitive process
  - Opportunity for NOAA to take stock and make adjustments if necessary

OVERVIEW

- **Objective**
  - Accelerate NWP Research to Operations (R2O) transition
- **Approach**
  - O2R
    - Make operational NWP systems available to research community
      - Code repositories, helpdesk, tutorials, etc
  - Test and Evaluation (T&E) of emerging research innovations
  - Engage community
    - Workshops, Visitor Program, etc
- **Task areas**
  - Mesoscale modeling (WRF ARW, NMMe, NMMb)
  - Data assimilation (GSI)
  - Hurricane forecasting (HWRF)
  - Ensemble forecasting (SREF)
  - Verification (MET)
- **Links with other NOAA Testbeds & programs**
  - HMT, HWT, HFIP

ACCOMPLISHMENTS

- **O2R** – Major accomplishments
  - Code repositories
    - WRF, GSI, HWRF, MET for community use; SREF for internal T&E
  - Helpdesks, workshops, tutorials, etc
  - Testing environment functionally similar to EMC's

- **R2O** – Significant T&E work
  - Reference configurations
  - Improvements to operational systems
  - Other experiments informing decisions regarding operational systems

Mesoscale Modeling

Jamie Wolf

Mesoscale Modeling
AOP 2012 Activities

Activity Description	Status
WRF-based community code maintenance and support: Repository maintenance, email support, code releases, tutorial	Ongoing
Physics interoperability for WRF-based system	In progress
Enhancement of NEMS-based code management: Technical discussions, friendly user release, FSOE for internal T&E	In progress
Establish a Mesoscale Model Evaluation Testbed (MMET)*: Define process for R2O transition, provide datasets and baseline results for cases of interest	Complete
Continue to conduct extensive T&E through comprehensive research innovation inter-comparisons and Reference Configuration designation: AFWA: WRF version difference and LIS input data set impact* NOAA: Surface drag parameterization schemes impact on a High Resolution Window WRF-ARW baseline configuration	AFWA – Complete NOAA – In progress

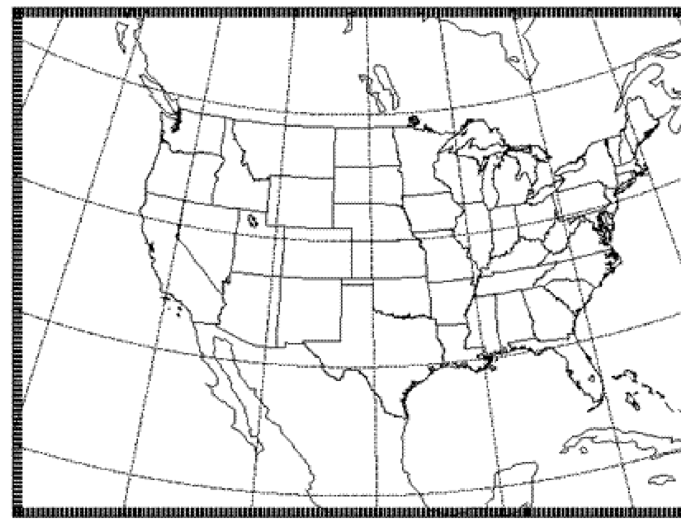
Key Accomplishments

Inter-comparison Testing and Evaluation

MMET

WRF Testing and Evaluation (T&E)

- End-to-end system: WPS, WRFDA, WRF, UPP, and MET
- Test Period: 1 July 2011 – 29 June 2012
- Retrospective forecasts: 48-h warm start forecasts initialized every 36 h w/ DA
- Domain: 15-km CONUS grid
- Evaluation:
  - Surface and Upper Air ((BC)RMSE, bias)
    - Temperature, Dew Point Temperature, Winds
  - Precipitation (GSS, frequency bias)
    - 3-h and 24-h accumulations
  - GO Index
  - Statistical Significance Assessment
    - Compute confidence intervals (CI) at the 99% level
    - Apply pair-wise difference methodology
    - Compute statistical significance (SS) and practical significance (PS)



WRF Inter-comparison T&E

- Functionally similar operational environment testing
  - WRF Data Assimilation and 6-hr warm start

	Current AFWA Op Configuration
Microphysics	WRF Single-Moment 5 scheme
Radiation SW and LW	Dudhia/RRTM schemes
Surface Layer	Monin-Obukhov similarity theory
Land-Surface Model	Noah
Planetary Boundary Layer	Yonsei University scheme
Convection	Kain-Fritsch scheme

  - WRFDAv3.3.1 + WRFv3.3.1 w/ LoBCs from LIS w/ Noahv2.7.1
  - WRFDAv3.4 + WRFv3.4 w/ LoBCs from LIS w/ Noahv2.7.1
  - WRFDAv3.4 + WRFv3.4 w/ LoBCs from LIS w/ Noahv3.3
- Evaluation included:
  - Impact assessment of WRF system version
  - Performance assessment of the LIS input data set

WRF v3.3.1 – v3.4 Results

- SS (light shading) and **PS (dark shading)** pair-wise differences for the annual aggregation of *surface temp, dew point and wind BCRMSE* and *bias* aggregated over the full set of cases and the entire integration domain

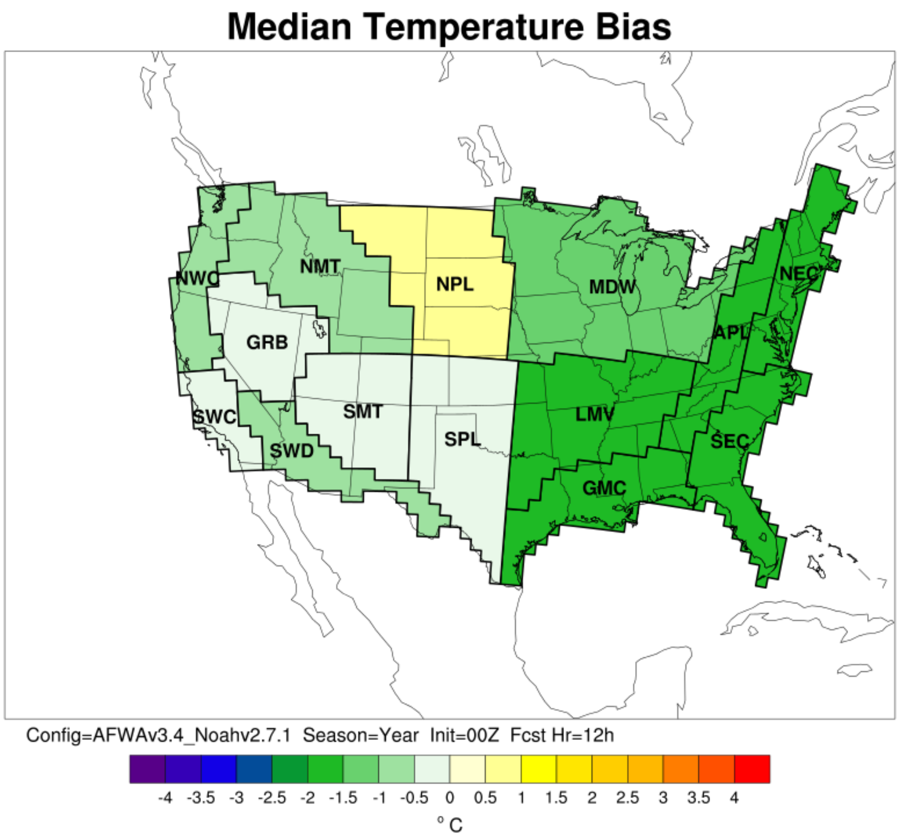
Annual			f03	f06	f09	f12	f15	f18	f21	f24	f27	f30	f33	f36	f39	f42	f45
BCRMSE	00 UTC Inits	Temperature	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	–	–	–	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	–	–
		Dew Point Temperature	v3.3.1	v3.3.1	v3.3.1	–	–	v3.3.1	v3.3.1	v3.3.1	–	–	–	–	–	–	v3.3.1
		Wind	v3.3.1	v3.3.1	v3.3.1	v3.3.1	–	–	–	v3.3.1	v3.3.1	v3.3.1	–	v3.3.1	–	–	–
	12 UTC Inits	Temperature	v3.3.1	–	–	–	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	–	–	–	v3.3.1	v3.3.1	v3.3.1
		Dew Point Temperature	v3.3.1	v3.3.1	–	v3.3.1	–	–	–	–	–	–	v3.3.1	v3.3.1	–	–	–
		Wind	–	v3.3.1	v3.3.1	–	v3.3.1	v3.3.1	v3.3.1	v3.3.1	–	–	–	–	v3.3.1	v3.3.1	–
Bias	00 UTC Inits	Temperature	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1
		Dew Point Temperature	–	v3.3.1	v3.3.1	v3.3.1	–	v3.3.1	v3.3.1	v3.4	v3.3.1	v3.3.1	v3.3.1	v3.3.1	–	v3.3.1	v3.3.1
		Wind	v3.4	v3.4	v3.4	v3.4	v3.4	–	–	v3.4	v3.4	v3.4	v3.4	v3.4	v3.4	–	–
	12 UTC Inits	Temperature	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1	v3.3.1
		Dew Point Temperature	–	v3.3.1	v3.3.1	–	v3.4	v3.3.1	v3.3.1	v3.3.1	–	v3.3.1	v3.3.1	v3.4	v3.3.1	v3.3.1	v3.3.1
		Wind	–	–	–	v3.4	v3.4	v3.4	v3.4	v3.4	–	–	–	v3.4	v3.4	v3.4	v3.4

Annual		Temperature				Dew Point Temperature				Wind			
		f12	f24	f36	f48	f12	f24	f36	f48	f12	f24	f36	f48
BCRMSE	850	–	–	–	–	v3.3.1	–	v3.3.1	v3.3.1	–	v3.3.1	–	–
	700	–	–	–	–	–	–	v3.3.1	–	–	v3.3.1	–	–
	500	v3.3.1	–	v3.3.1	v3.3.1	–	v3.3.1	–	–	v3.3.1	–	v3.3.1	v3.3.1
	400	–	–	–	–					–	–	v3.3.1	v3.3.1
	300	v3.3.1	–	–	–					v3.3.1	v3.3.1	v3.3.1	v3.3.1
	200	–	–	–	v3.3.1					–	–	v3.3.1	–
	150	–	–	v3.3.1	v3.3.1					v3.3.1	–	–	v3.3.1
	100	v3.3.1	v3.3.1	v3.3.1	–					v3.3.1	v3.3.1	v3.3.1	v3.3.1
Bias	850	v3.3.1	v3.3.1	–	–	v3.4	v3.4	v3.4	v3.4	v3.4	v3.4	–	v3.4
	700	v3.3.1	v3.3.1	v3.3.1	–	v3.4	v3.4	–	–	v3.4	v3.4	v3.4	v3.4
	500	v3.4	v3.4	v3.4	–	–	–	v3.4	–	v3.4	v3.4	v3.4	v3.4
	400	v3.4	v3.4	v3.4	–					v3.4	v3.4	v3.4	v3.4
	300	v3.4	v3.4	v3.4	v3.4					v3.4	v3.4	v3.4	v3.4
	200	v3.4	v3.4	v3.4	v3.4					v3.4	v3.4	v3.4	v3.4
	150	v3.3.1	v3.3.1	v3.3.1 *	v3.3.1 *					v3.4	v3.3.1	v3.3.1	v3.3.1
	100	v3.4 *	v3.4 *	v3.4 *	v3.4 *					v3.3.1	v3.3.1	v3.3.1	v3.3.1

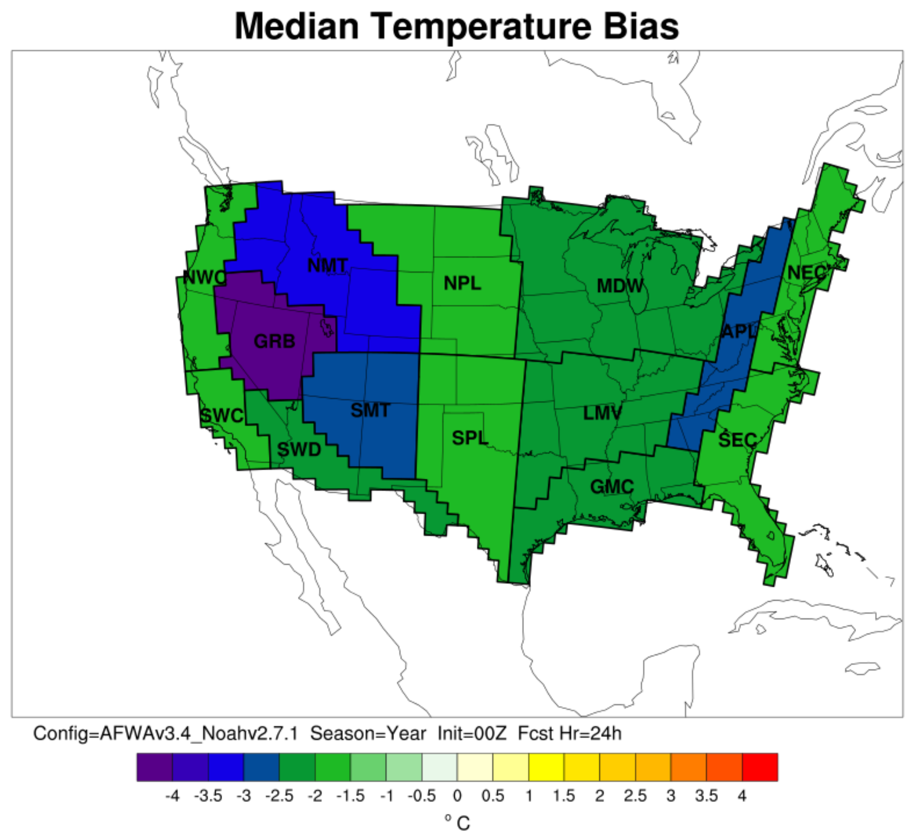
Regional Temperature Bias Verification

WRF v3.4.1 w/ Noah v2.7.1

00 UTC 12h forecast



00 UTC 24h forecast



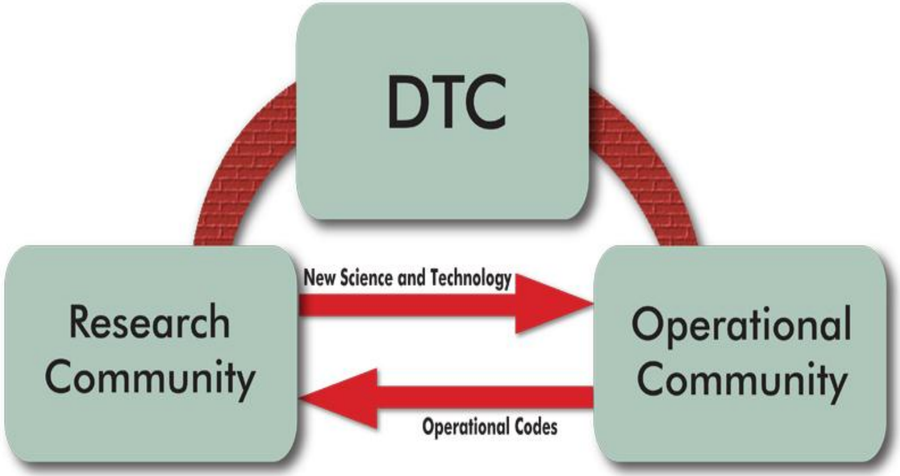
Key Accomplishments

Inter-comparison Testing and Evaluation

MMET

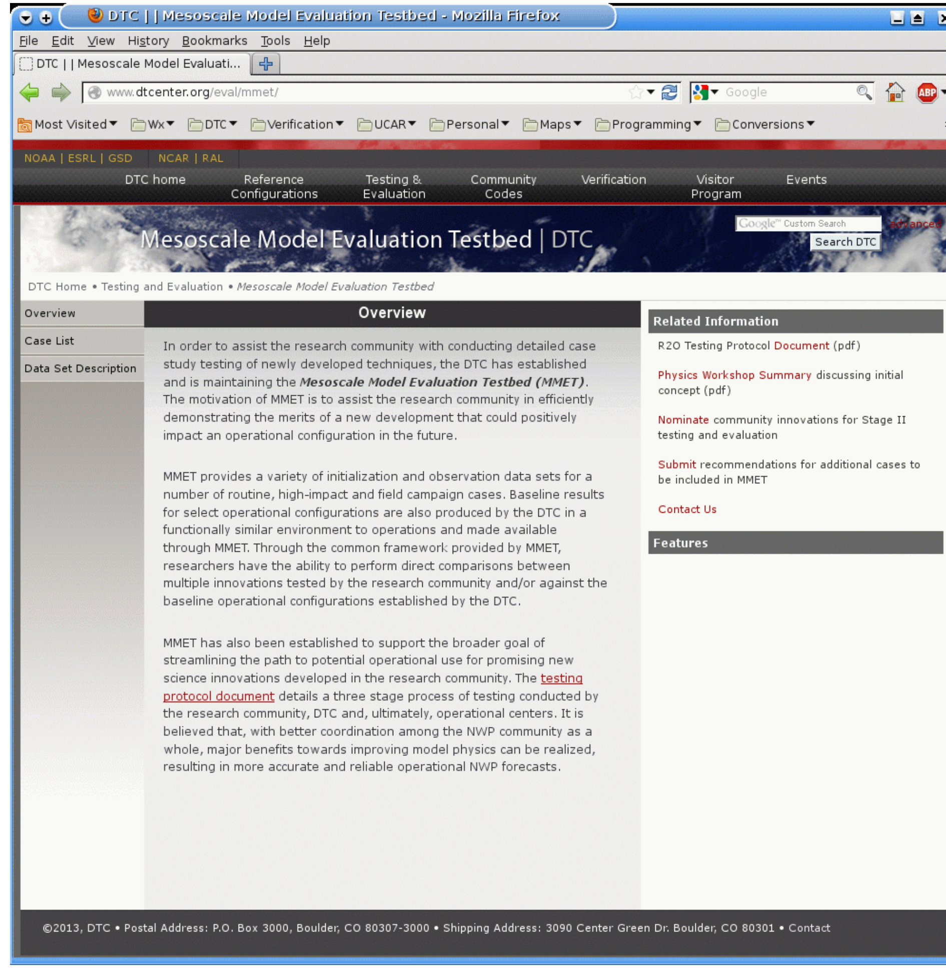
Testing Protocol Motivation

- Wide range of NWP science innovations under development in the research community
- Testing protocol imperative to advance new innovations through the research to operations (R2O) process *efficiently* and *effectively*.
  - Three stage process:
    1) Proving ground for research community
    2) Comprehensive T&E performed by the DTC
    3) Pre-implementation testing at Operational Centers



Mesoscale Model Evaluation Testbed (MMET)

- What: Mechanism to *assist* research community *with initial stage of testing* to efficiently demonstrate the merits of a new development
  - Provide model input and observational datasets to utilize for testing
  - Establish and publicize baseline results for select operational models
  - Provide a common framework for testing; allow for direct comparisons
- Where: Hosted by the DTC; served through **R**epository for **A**rchiving, **M**anaging and **A**ccessing **D**iverse **DA**ta (RAMADDA)



www.dtcenter.org/eval/mmet

Hurricane

Ligia R. Bernardet

External collaborators:
NOAA Environmental Modeling Center
NOAA Geophysical Fluid Dynamics Laboratory
NOAA Atlantic Oceanographic and Meteorological Laboratory
NCAR Mesoscale and Microscale Meteorology Division
University of Rhode Island
University of California – Los Angeles
Florida State University

Hurricane AOP 2012 Activities

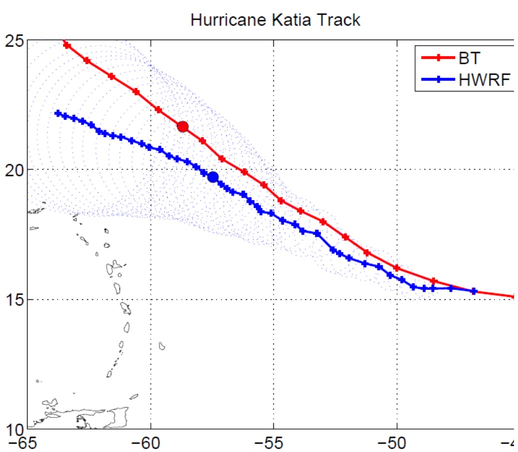
Activity Description	Status
Software systems & community support activities	
HWRF repository maintenance, public release and user support	Ongoing
HWRF interoperability – Thompson microphysics	In progress
HWRF FSOE to match 2012 operational	Competed
T&E activities	
HWRF 2012 operational Reference Configuration	Completed
T&E FSOE: HWRF cumulus sensitivity	Completed
T&E FSOE: HWRF atmos-ocean fluxes	**Completed**
Sensitivity experiments: Thompson microphysics in HWRF	Current– will complete in Feb
Diagnostics of large scale environment in HWRF	Completed

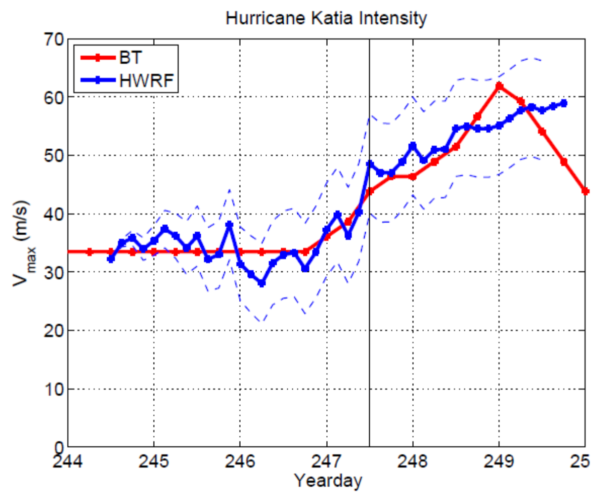
POM Flux Test

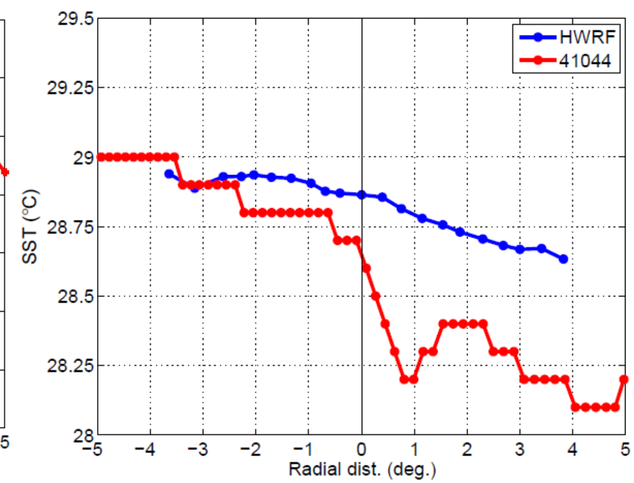
Background

HRD (Uhlhorn and Cione) compared HWRF retro forecasts for 2011 against buoys and showed that HWRF ocean does not respond (=does not cool as much as obs) when storm goes by

Example: Katia 09/01/11 init12 UTC and buoy passage 9/4 12 UTCZ

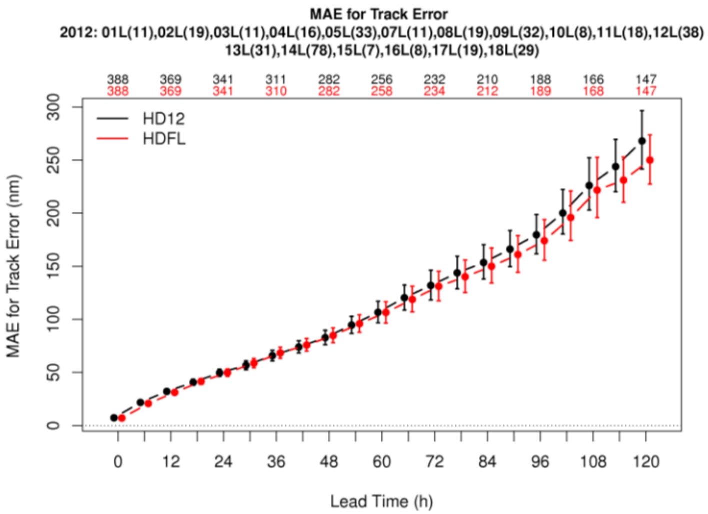


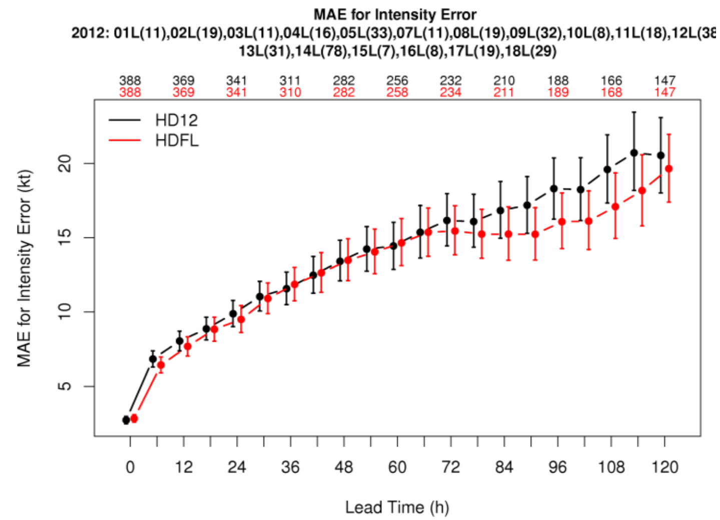


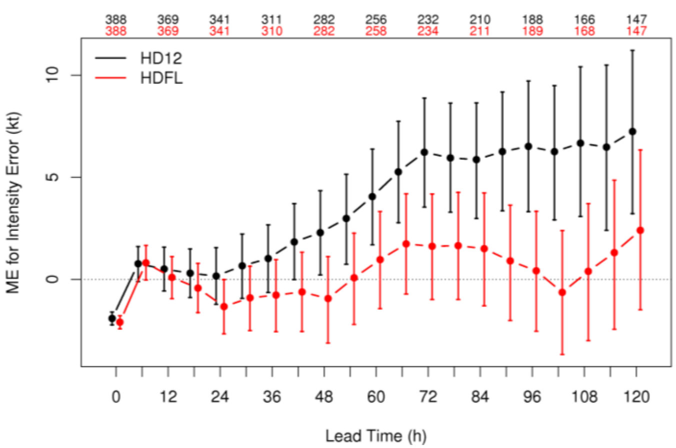


- Fluxes from HWRF atmosphere to ocean are truncated in POM (75%)
- DTC ran 2012 season: control HD12 (75% fluxes) and modified HDFL (100%)

Atlantic track and intensity

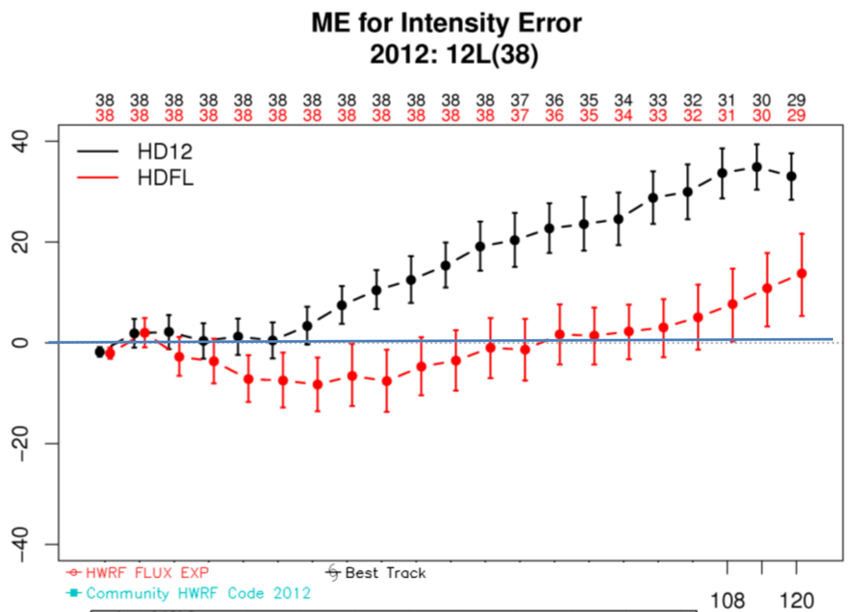


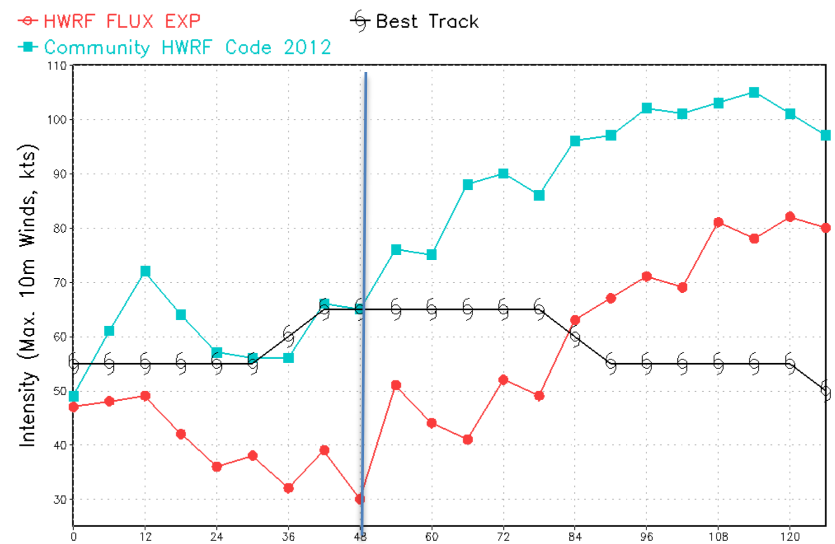


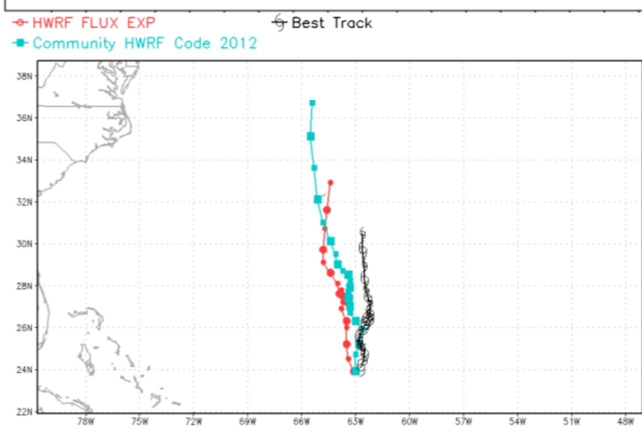


Track ME: HD12 and HDFL very similar
Int MAE: HDFL SS better at 3 lead times
Int bias: HD12 lowers intensity and helps overintensification at long lead times
Hurricane Leslie (12L) is the storm with largest impact (large and slow)
Pacific impact is much smaller (POM 1D)

Leslie bias and 09/04 00Z case



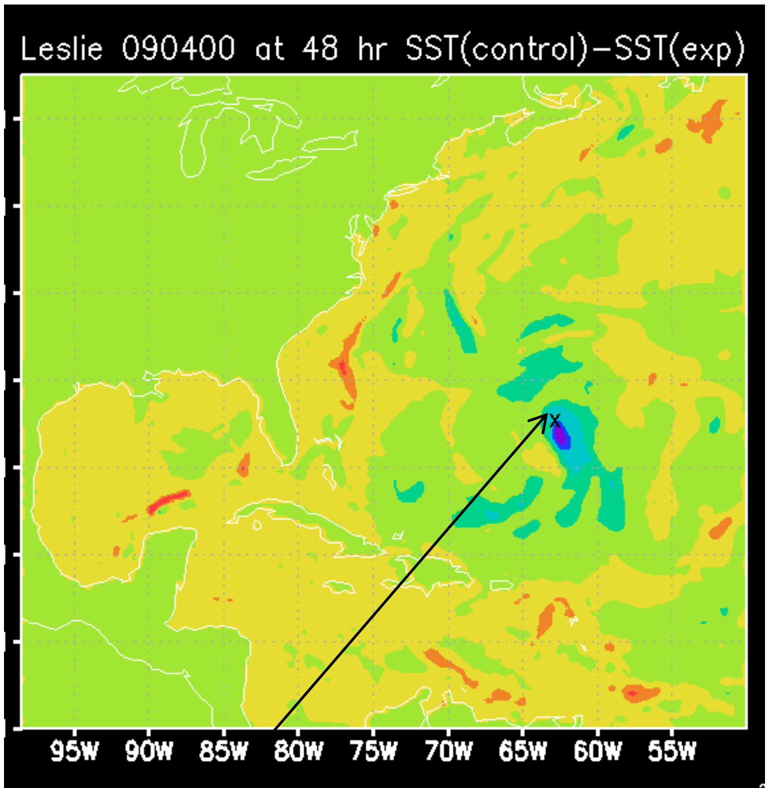




- HD12 and HDFL tracks are similar
- HDFL reduces intensity (as expected).
- Is it because of low SST under storm?

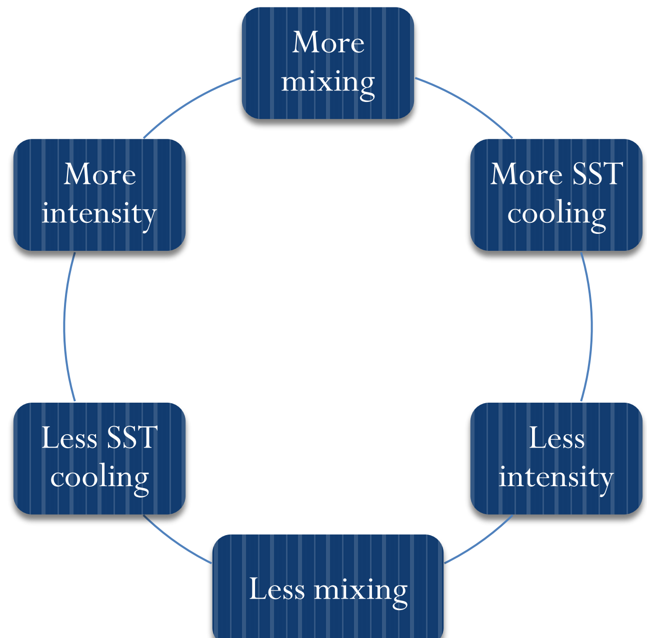
Leslie bias and 09/04 00Z case

48-h SST control – flux exp



X = storm center

At 48 h, control has cooler SST than flux exp (contrary to linear interpretation)



Data Assimilation

Hui Shao

Acknowledgements:
HFIP, EMC, Brian Etherton, Ligia Bernardet

Mechanism for DTC Data Assimilation T&E

Operational GSI implementation and parallel test runs. Focus on evaluating the overall performance of GSI.

- Benchmark
- Parallel run config
- Archived data /background for retro runs

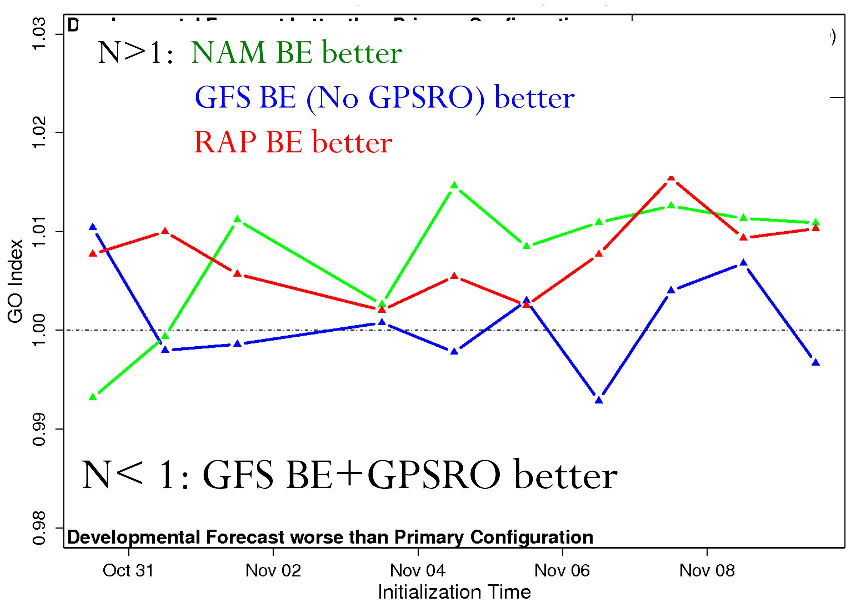
DTC real-time & retrospective GSI runs using functionally-similar operational environment: Focus on testing incremental changes.

- **Real-time**: "sync" testbed, uncover the issues
- **Short-term retrospective**: test individual changes, tackle the issues
- **Extensive retrospective**: impact study w/ SS, test research/developmental components

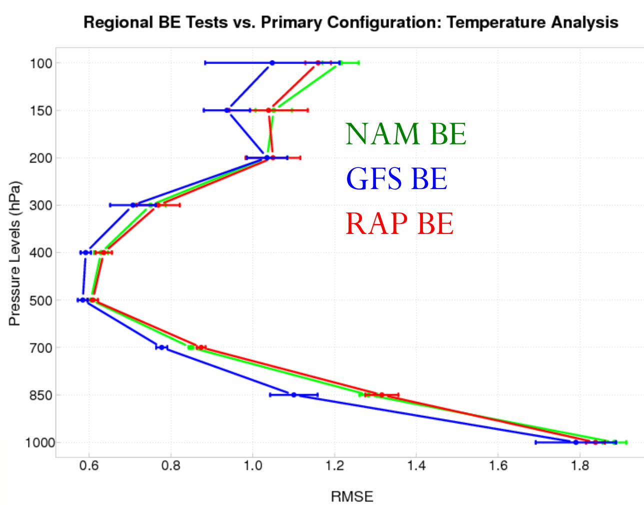
- Benchmark
- Developmental config (suggested from the DTC)

Pathway to "O"

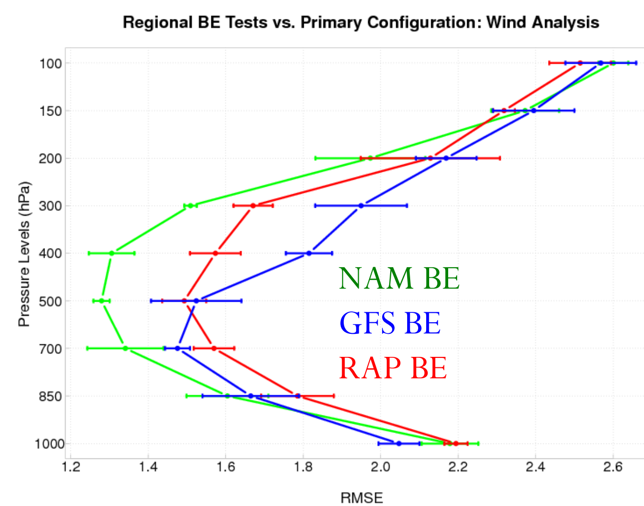
GSI Configuration T&E for Regional Applications



- ✓ NAM BE: Northern Hemisphere BE computed based on NAM forecasts.
- ✓ GFS BE: Global BE computed based on GFS forecasts.
- ✓ RAP BE: Global BE tuned for the RAP. combination of global/regional *(balance = GFS, Lengthscales/variance = NAM)*



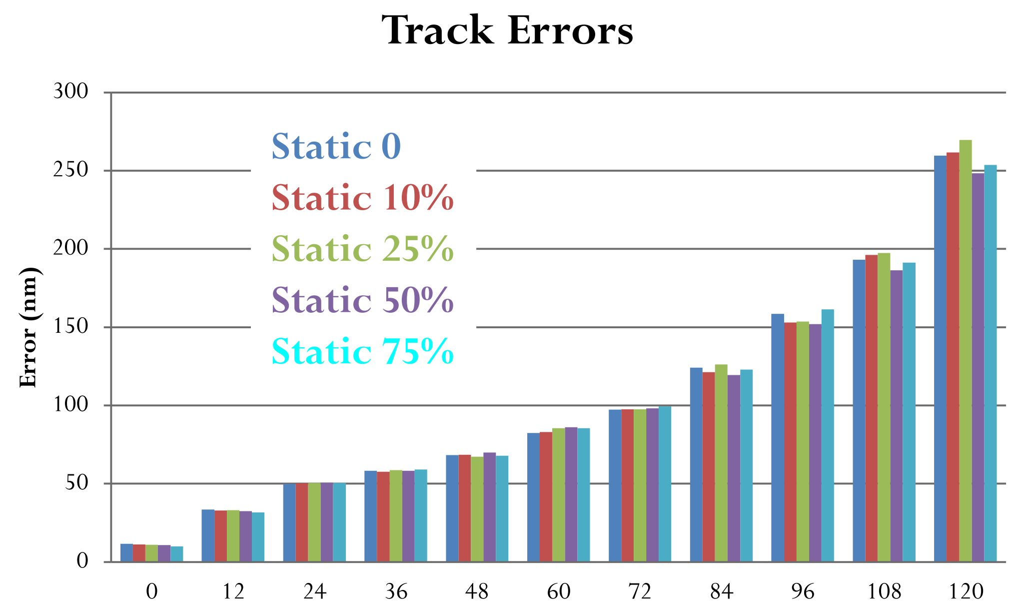
Temperature Analysis RMSE



Wind Analysis RMSE

GSI-Hybrid T&E for HWRF Applications

- Coordinated with HFIP GSI-hybrid tiger team members
- System examination and alternative configuration T&E:
  - Cross covariance
  - Cycling versus cold-start
  - Relative contributions of static background error (BE) and ensemble BE statistics



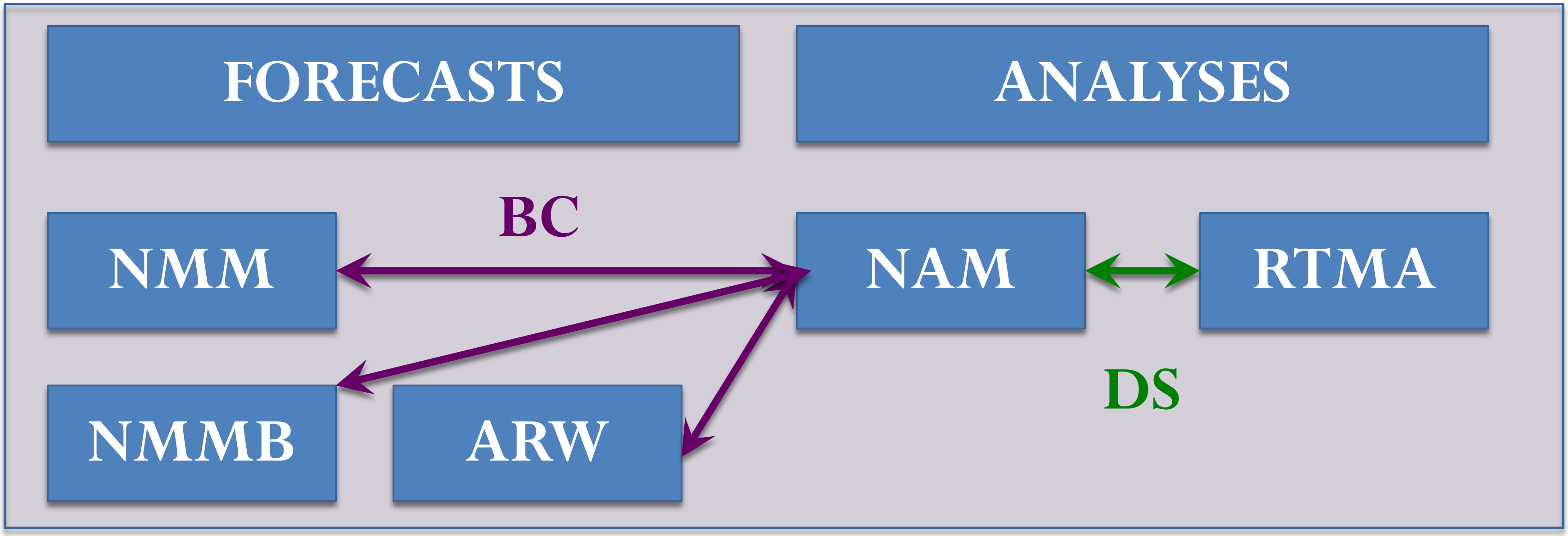
Ensemble Forecasting

Brian Etherton, Tara Jensen, Jun Du, Tara Jensen,
Isidora Jankov

Downscaling SREF

- SREF 2012 upgrade to 16 km resolution
  - Significant change from 30+ km
    - Still not enough for fine scale features needed for NDFD
- Downscale 16 km SREF to 5 km NDFD
  - Apply and test North American Ensemble Forecast System (NAEFS) downscaling algorithm

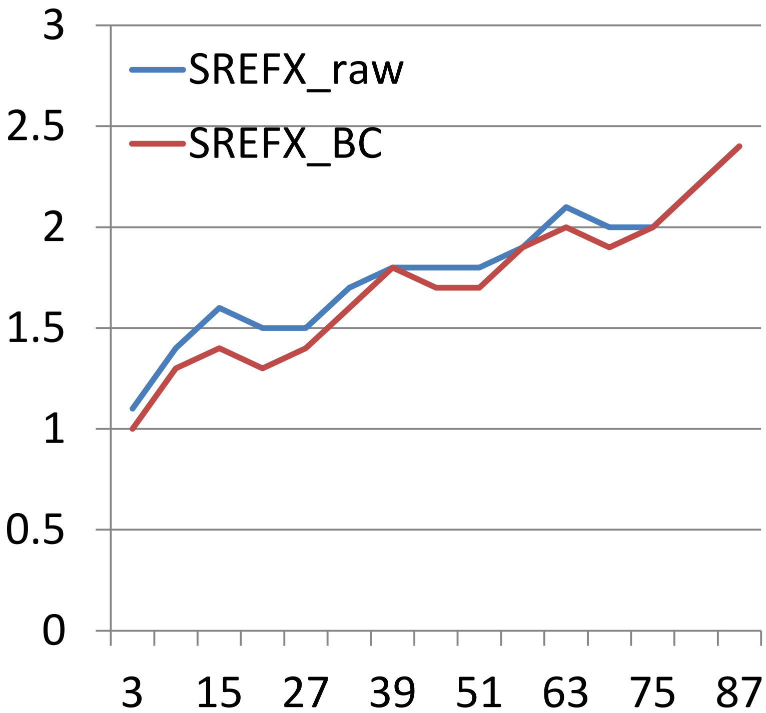
Bias Correction and Downscaling



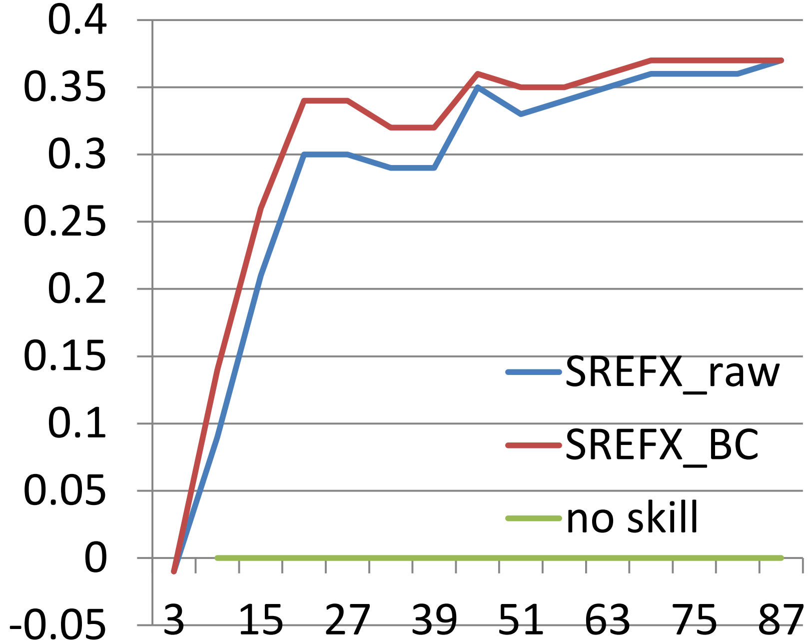
- Bias Correction – NAEFS, also in SREF operations
  - Take mean forecast of each model core (ARW, NMM, etc.) sub-ensemble of SREF
  - Compare them to NAM analysis valid at the same time
- **Downscaling – NAEFS – adapted and tested for possible use in SREF**
  - Compare RTMA analysis (5 km) with NAM analysis interpolated to same NDFD grid
    - 10m wind,2m temperature, humidity – analyses valid at same time
- Recursive averaging to estimate biases (~30 day mean) & downscaling (~5 days)
- Bias corrected and downscaled fields for each member

Bias Correction at EMC

T2m Ens mean fcst: **RMSE**



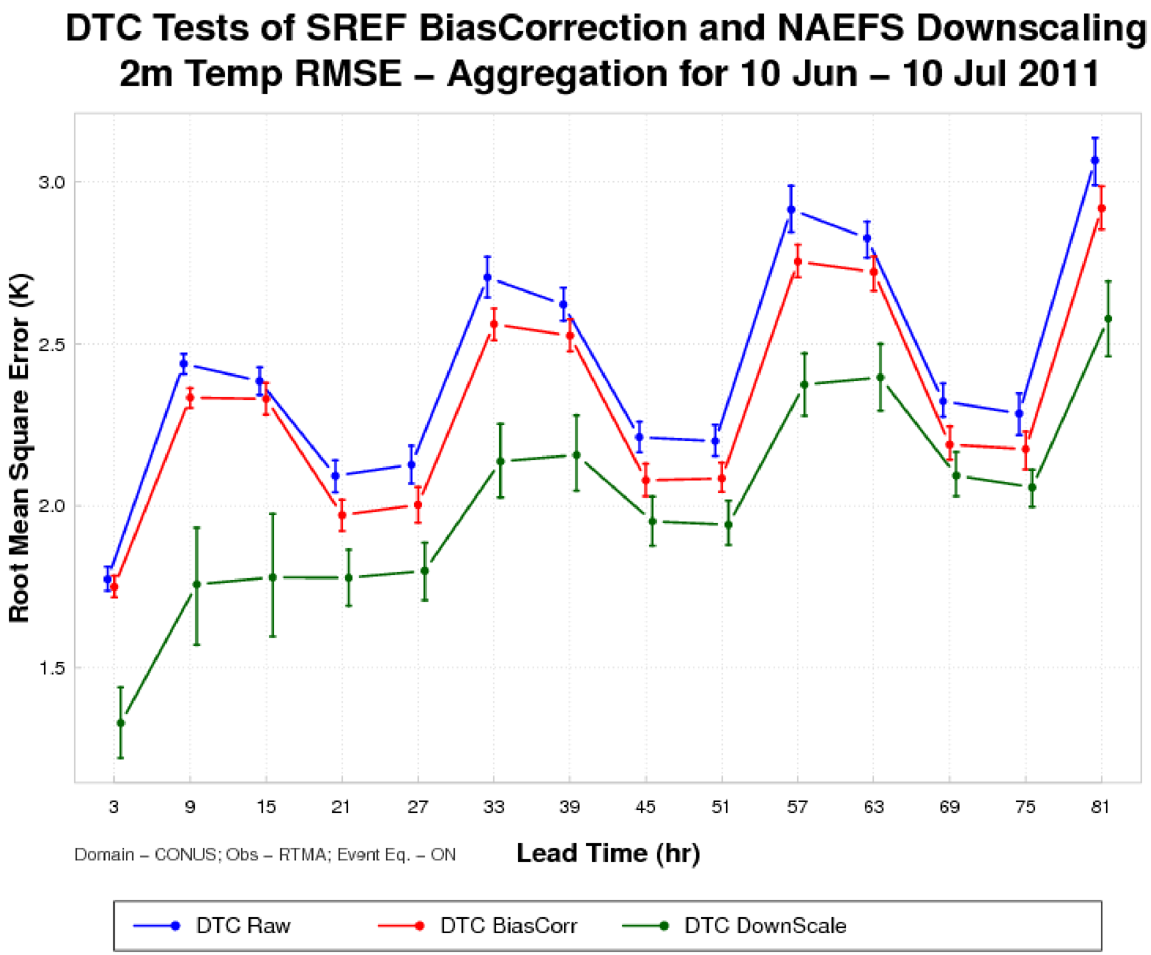
T2m Prob fcst: **RPSS** (12km NAM as ref)



Raw SREFx vs Bias corrected SREFx (Nov. 10 – 30, 2011, against NDAS)

Courtesy of Jun Du, NOAA/EMC

Testing/Evaluation at DTC - Results



ARW and NMM members of SREF 2011 – 0900 UTC Initialization Compared to RTMA Analyses

Downscaling much reduces the error in the bias corrected 2m temperature forecasts

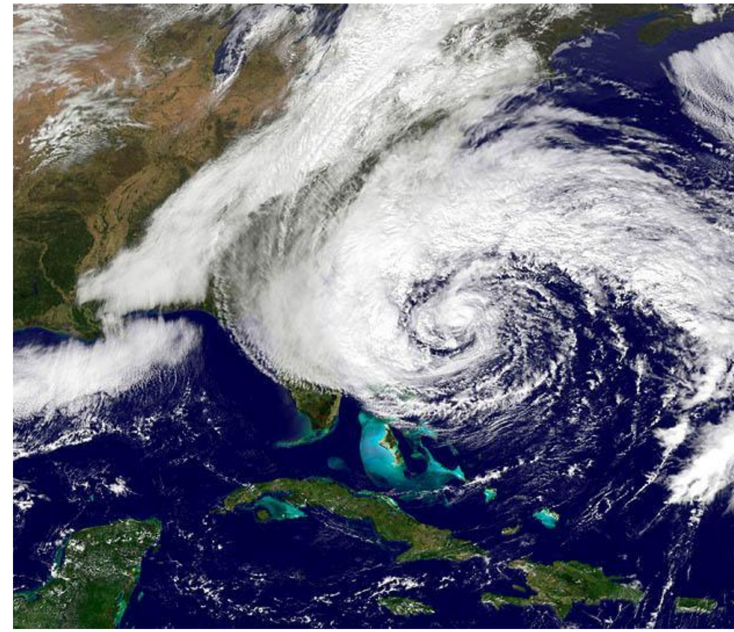
Verification

Tressa Flower

DTC Verification Accomplishments

Software Development

- METTC
- MADIS data support
- Ensemble spread skill
- GRIB2
- **Series analysis tool**



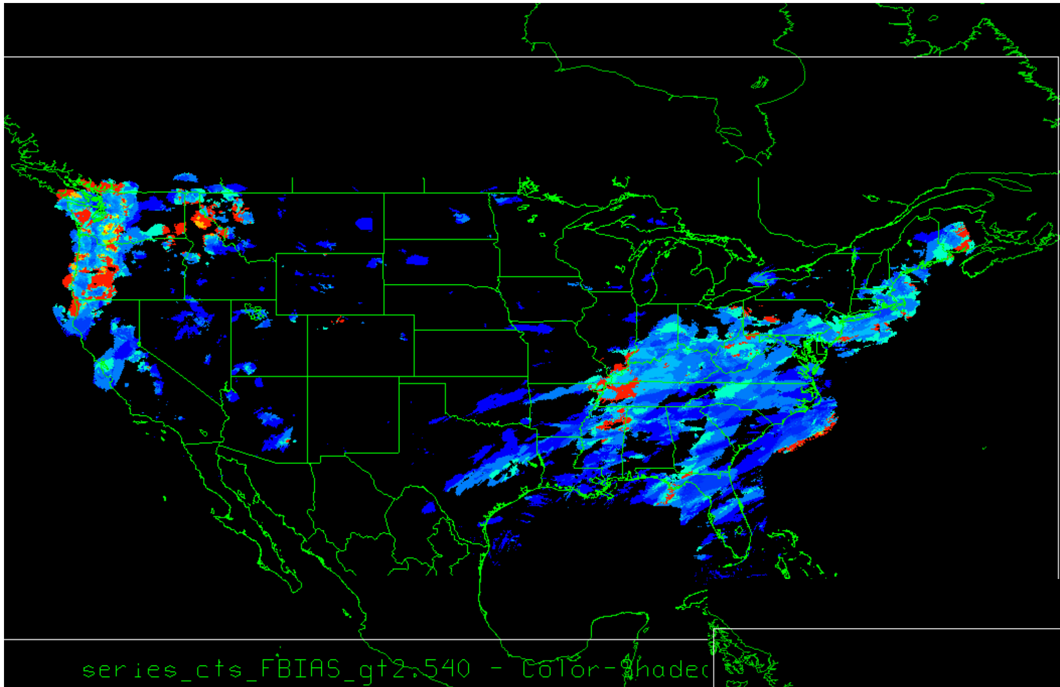
Testing and Evaluation support

- **HMT verification**
- MMET cases

Community support

MET
Model Evaluation Tools

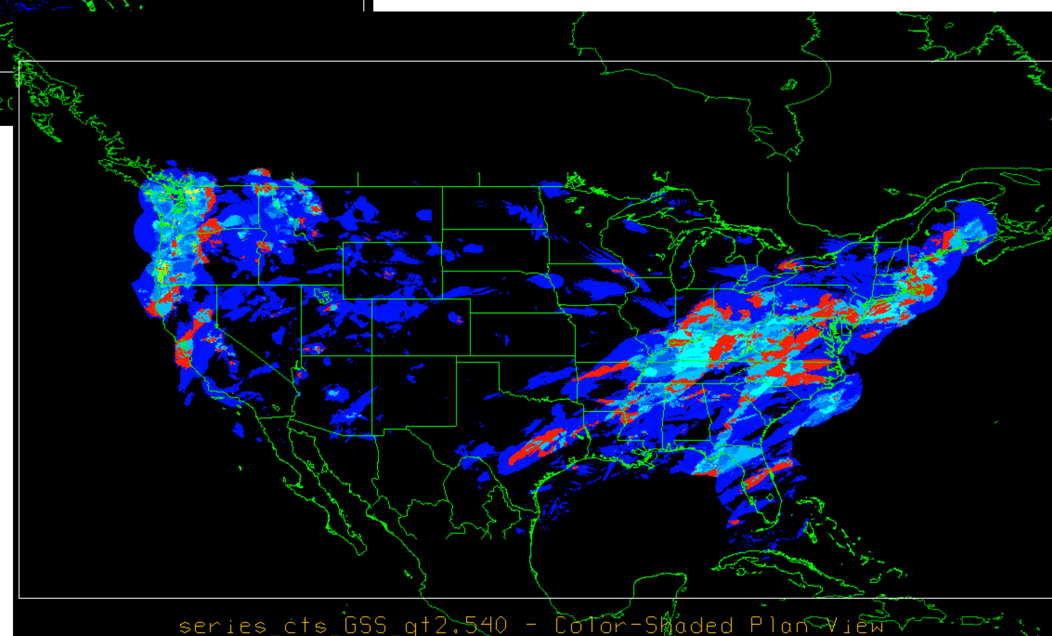
Series Analysis Tool Example



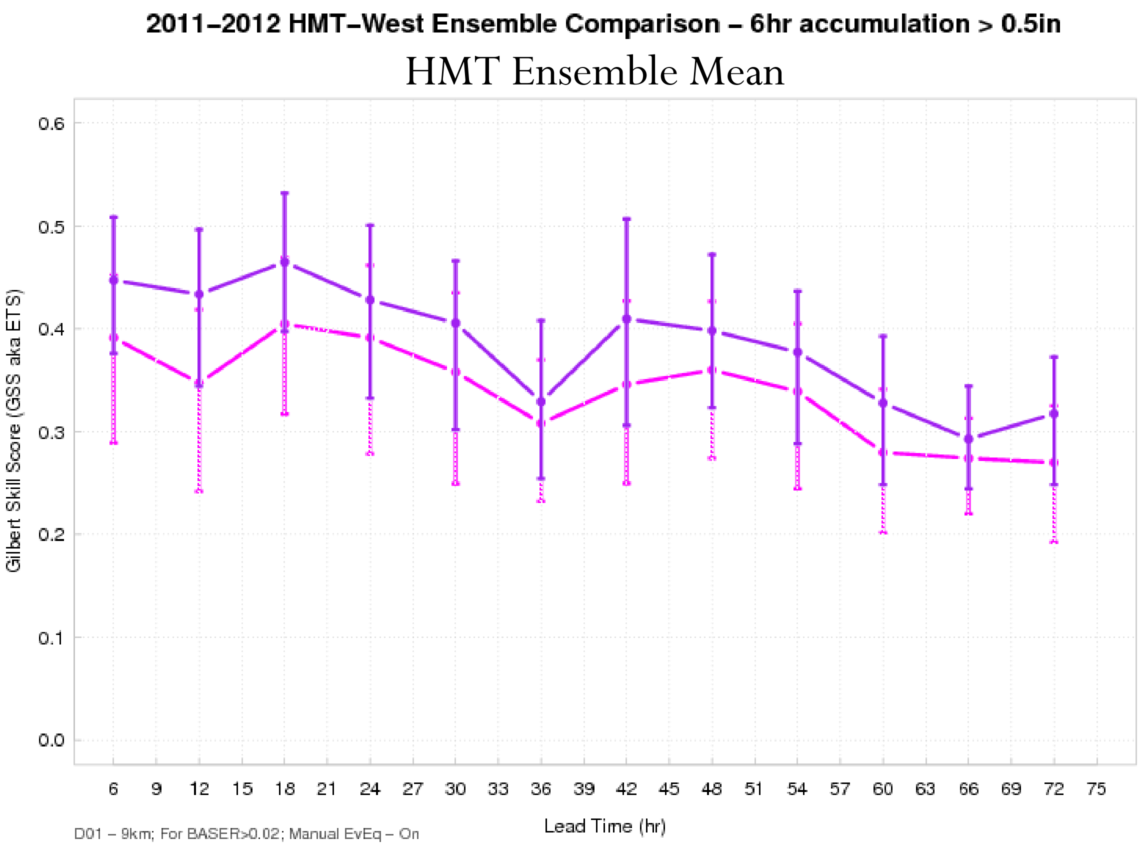
Statistics accumulated over time at each grid location

Gilbert Skill Score

Frequency Bias



Verification Support of HMT



No constraint

Base Rate > 0.02

Capability was added to METViewer:

User can constrain aggregation by observed relative frequency

Assess skill for events selected by threshold

Increases analysis speed and relevance

FUTURE OF DTC

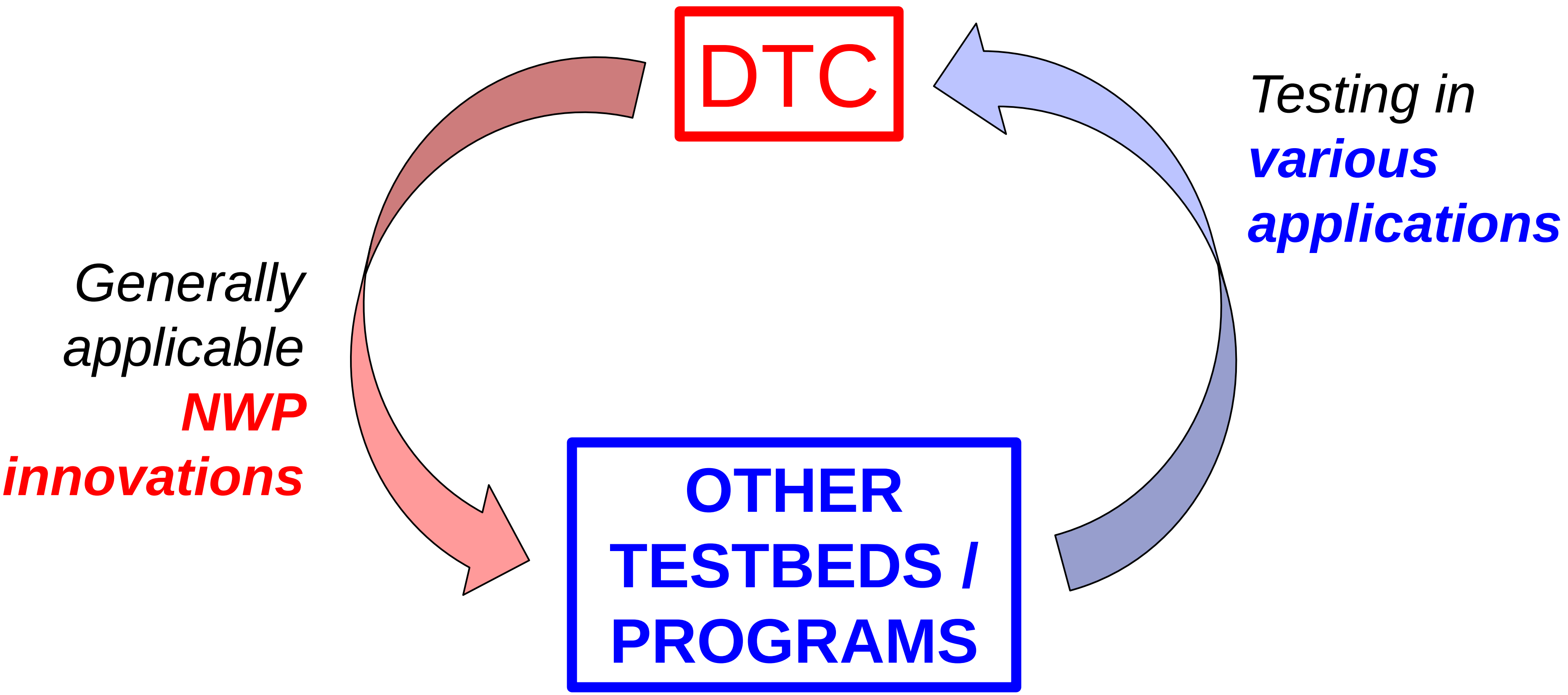
- Organization
- O2R & other support
- R2O
  - Current systems
  - Next generation systems

ORGANIZATIONAL CONSIDERATIONS

- Find best **organizational structure for DTC**
  - NOAA level
    - OAR and NWS collaboration
      - Define NOAA needs for new cooperative agreement
  - Interagency coordination
    - Leverage efforts by other agencies

- Strengthen **links with other NOAA testbeds and programs**
  - Ongoing collaboration with HFIP, HMT, HWT
  - Potential links with JCSDA, JHT, CTB, Satellite Proving Ground, others?
    - DTC / NWP testbed - results relevant for number of testbeds/programs
    - Other testbeds using NWP tools – application areas for DTC

DTC & OTHER TESTBEDS / PROGRAMS



SUPPORT FOR R2O

- Continue maintaining **unified DTC-EMC code repositories**
  - Necessary for T&E; success of DTC, resource intensive

- Create **new NWP Information Technology Environment** (NITE)
  - DTC created replica of operational environment for DTC T&E
    - Potentially inefficient approach; instead
  - Build modern interconnect NWP
    - Database, model launcher, display, verification, etc tools
    - To be shared & used by NCEP, DTC, their visitors
      - Systems like what ECMWF has

- Identify **support for academic PIs' R2O work**
  - Continue DTC Visitor Program
  - Engage NSF & other partners

HOW TO IMPROVE R2O?

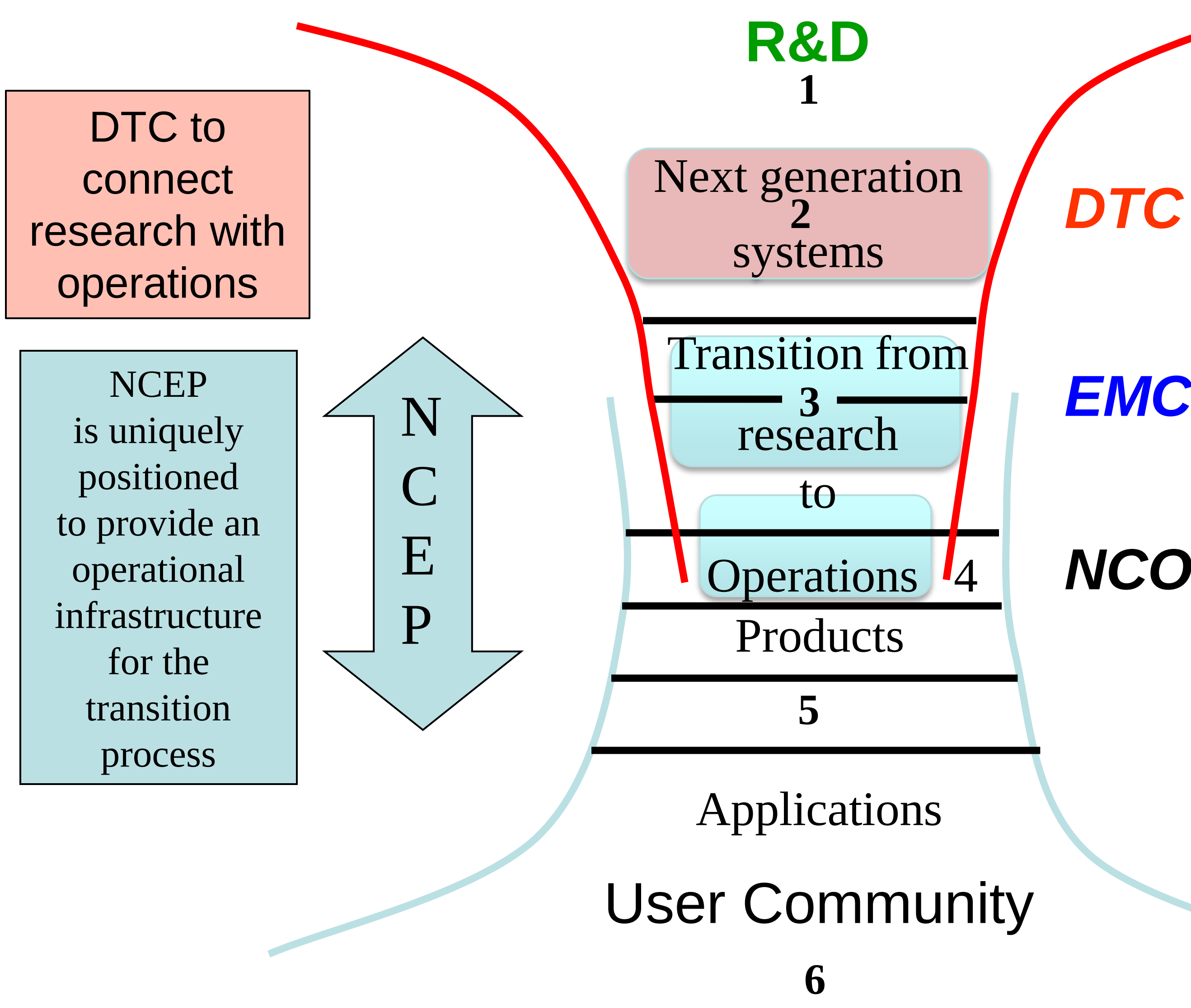
CURRENTLY OPERATIONAL SYSTEMS (1-2 year timeframe)

- Success with AFWA
  - Can be improved for NOAA
- T&E must be responsive to NCEP needs
- AOP must be aligned with NCEP plans

NEXT GENERATION SYSTEMS (3-5 year timeframe)

- Potentially large payoff
- Role of various partners
  - Academia — Basic research and method development
  - DTC — Building and testing prototype systems
  - EMC — Integrating into & testing in operational environment
- DTC must work with academia & EMC

DTC'S ROLE IN TRANSITION FUNNEL



1. Large "volume" of academic / agency lab R&D, **5 yrs**

2 Smaller set of R&D products suitable for operations. **3 yrs**

3. Systematic transition steps. **1 yr**

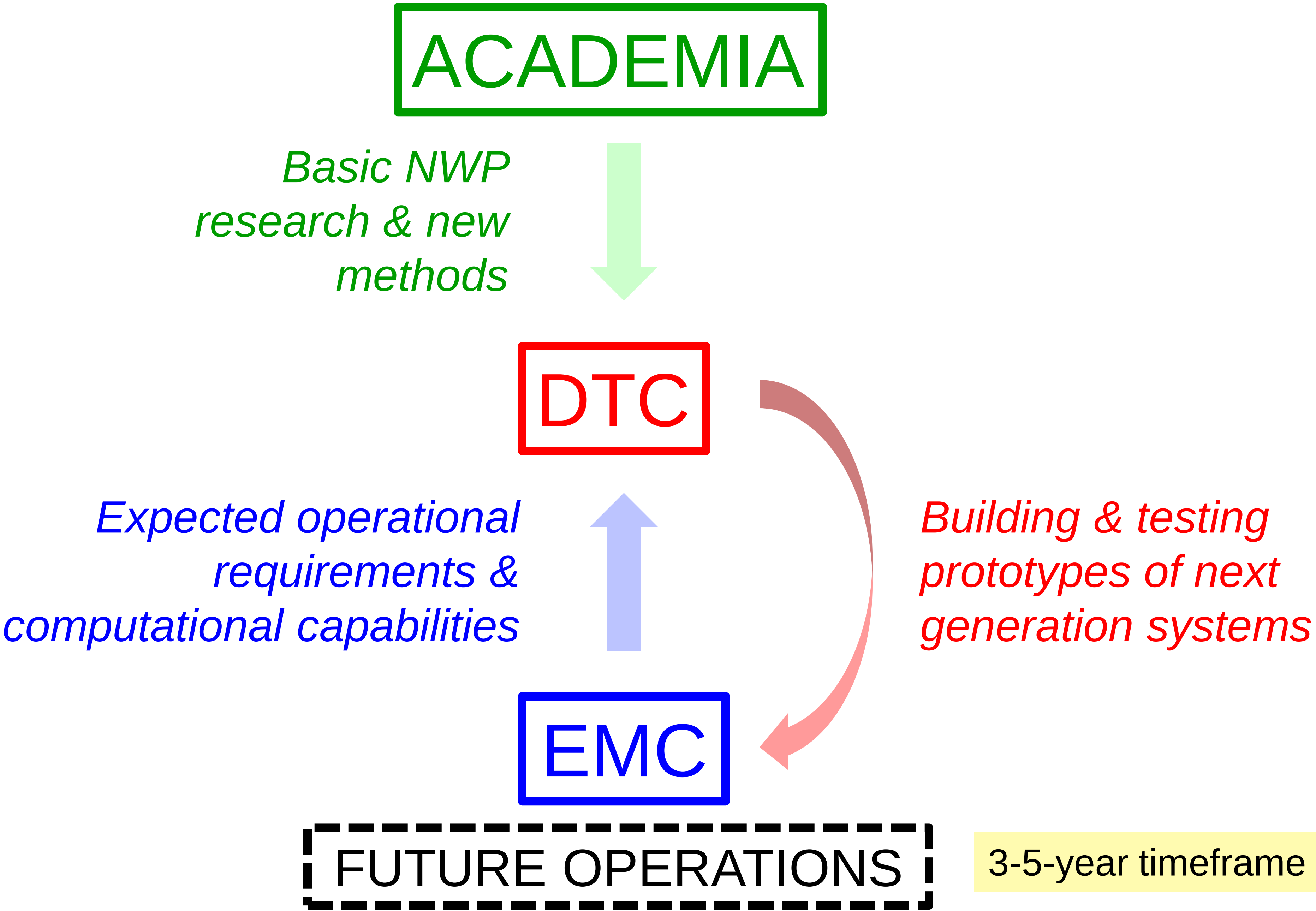
3. Operations - **Current**

3. New products can serve diverse and expanding user community.

6. Delivery to diverse USER community

After L. Uccellini & A. MacDonald

NEXT GENERATION NWP SYSTEMS



OUTLINE / SUMMARY

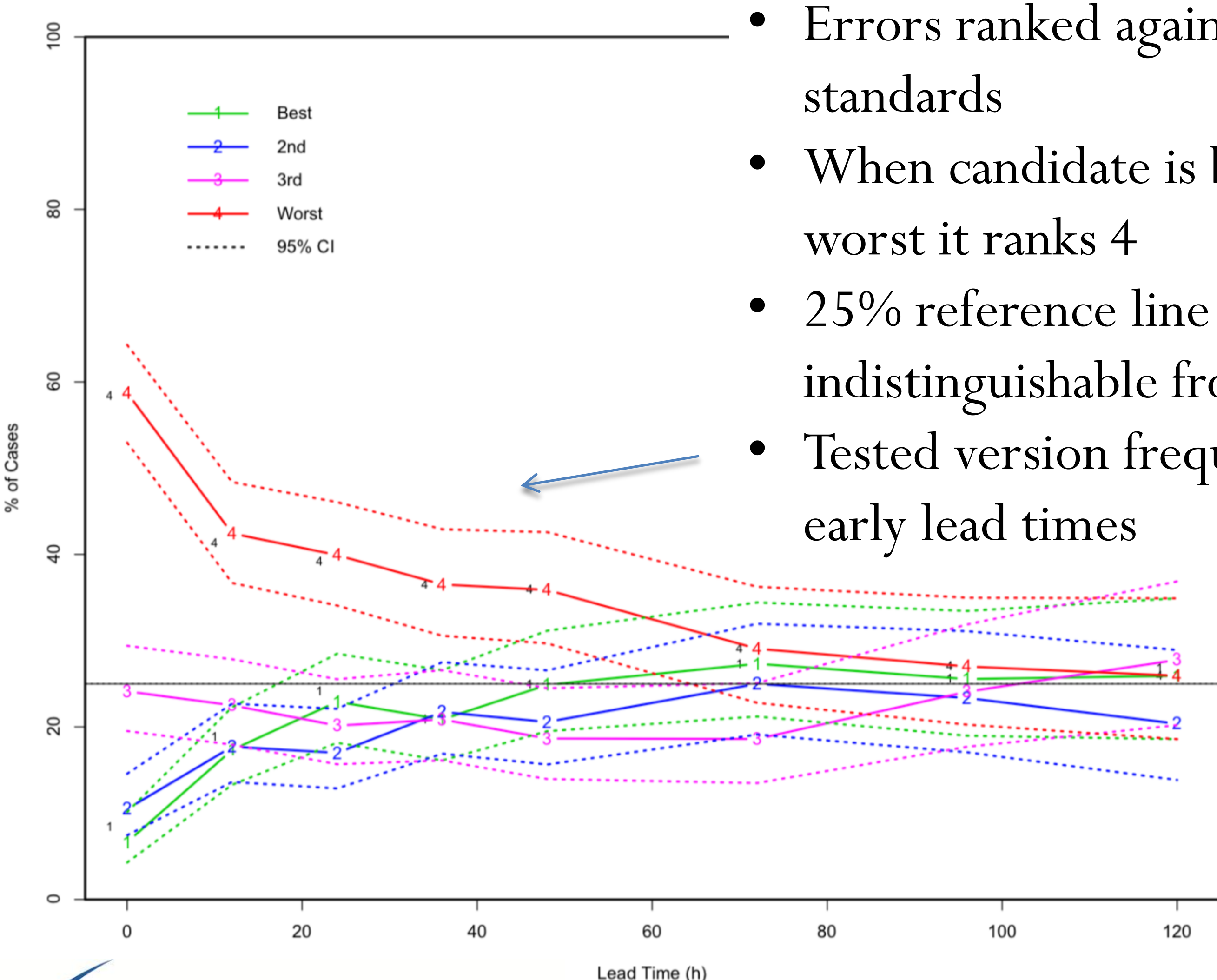
- **Overview**
  - Transition of research into operations
    - For Numerical Weather Prediction (NWP)

- **Research to Operations (R2O) testing**
  - WRF, HWRF, GSI, SREF, supported by
    - Operations to Research (O2R, e.g., code repositories)

- **Outlook**
  - Discussions on scope of DTC
    - Improve current & next generation NWP systems
    - New Cooperative Agreement
  - Build modern NWP IT Environment (NITE)
  - Strengthen collaboration with other NOAA testbeds & programs

BACKGROUND

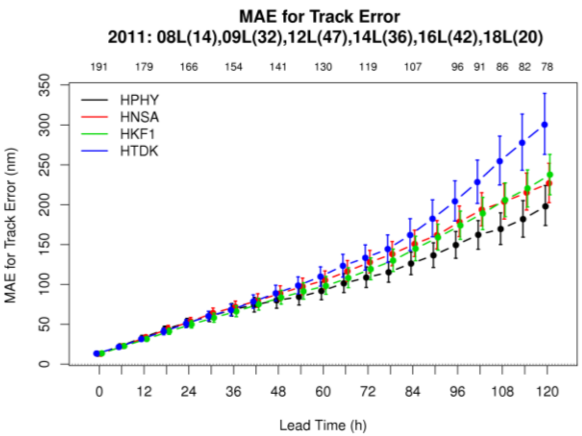
Track Error Rank of TC Model vs. 3 Operational Models

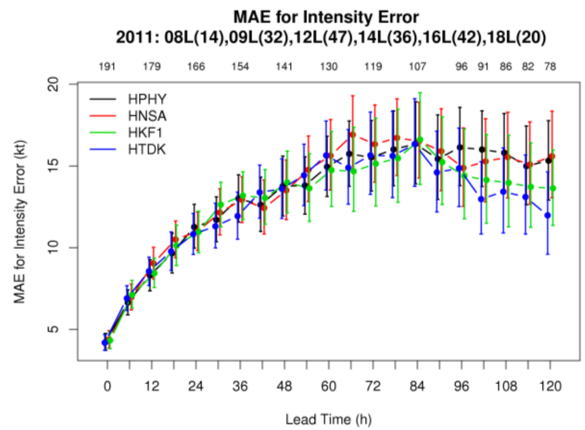


- Errors ranked against 3 operational standards
- When candidate is best it ranks 1 and when worst it ranks 4
- 25% reference line is for performance indistinguishable from the standards
- Tested version frequently ranked worst for early lead times

Cumulus sensitivity test

Test of HWRF sensitivity to cumulus schemes





Tested HWRF SAS, new SAS, Tiedtke, Kain-Fritsh

HWRF SAS performs best for track; differences in intensity have little statistical significance

Track

	12	24	36	48	60	72	84	96	108	120
HNSA										
HKF1										
HTDK										

Intens

	12	24	36	48	60	72	84	96	108	120
HNSA										
HKF1										
HTDK										

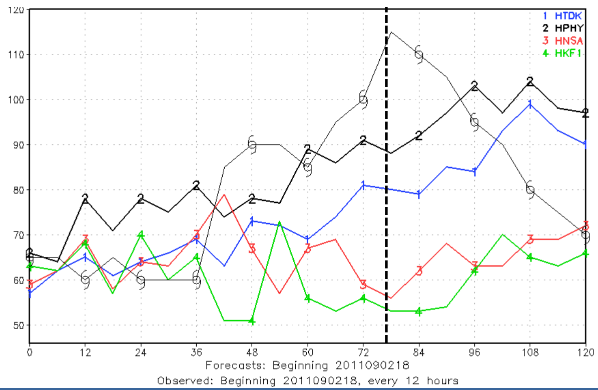
Statistical Significance 95%

Green= HWRF SAS better

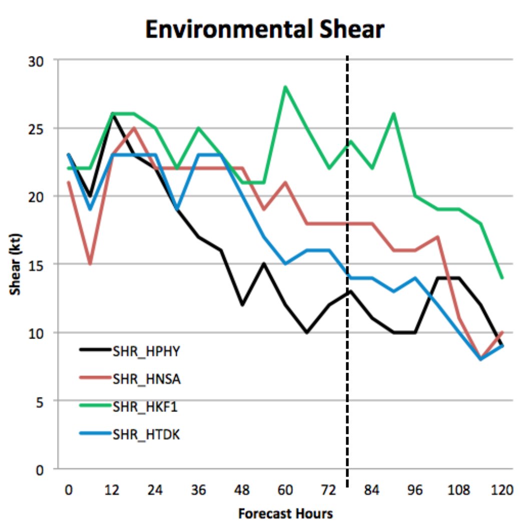
Red = HPHY SAS worse

Developmental Testbed Center

Case study: Katia init 09/02/11 18 Z,



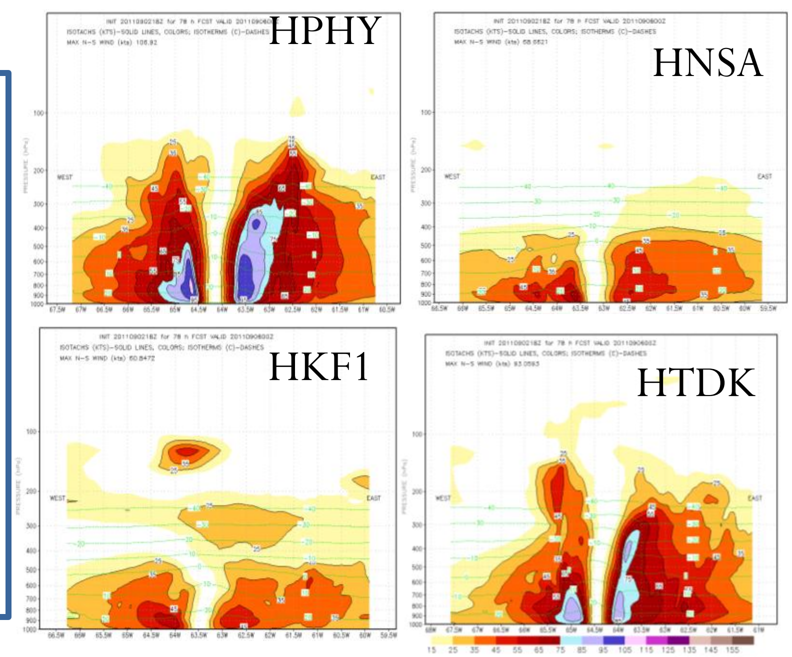
Tracks: similar
Intensity: different (HPHY, HTDK intensify)



SHIPS diagnostics of <u>shear</u>: initially similar, later different. Intensifiers have lower shear.

Highlights cumulus effects on and control on intensification

78-h forecast isotachs (E-W x-section)

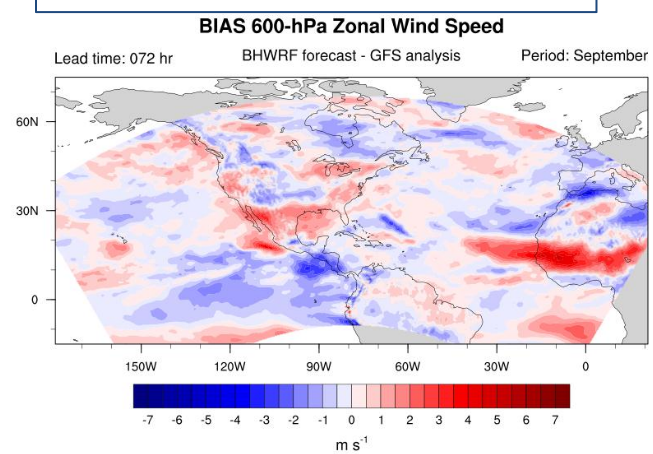


Large scale diagnostics

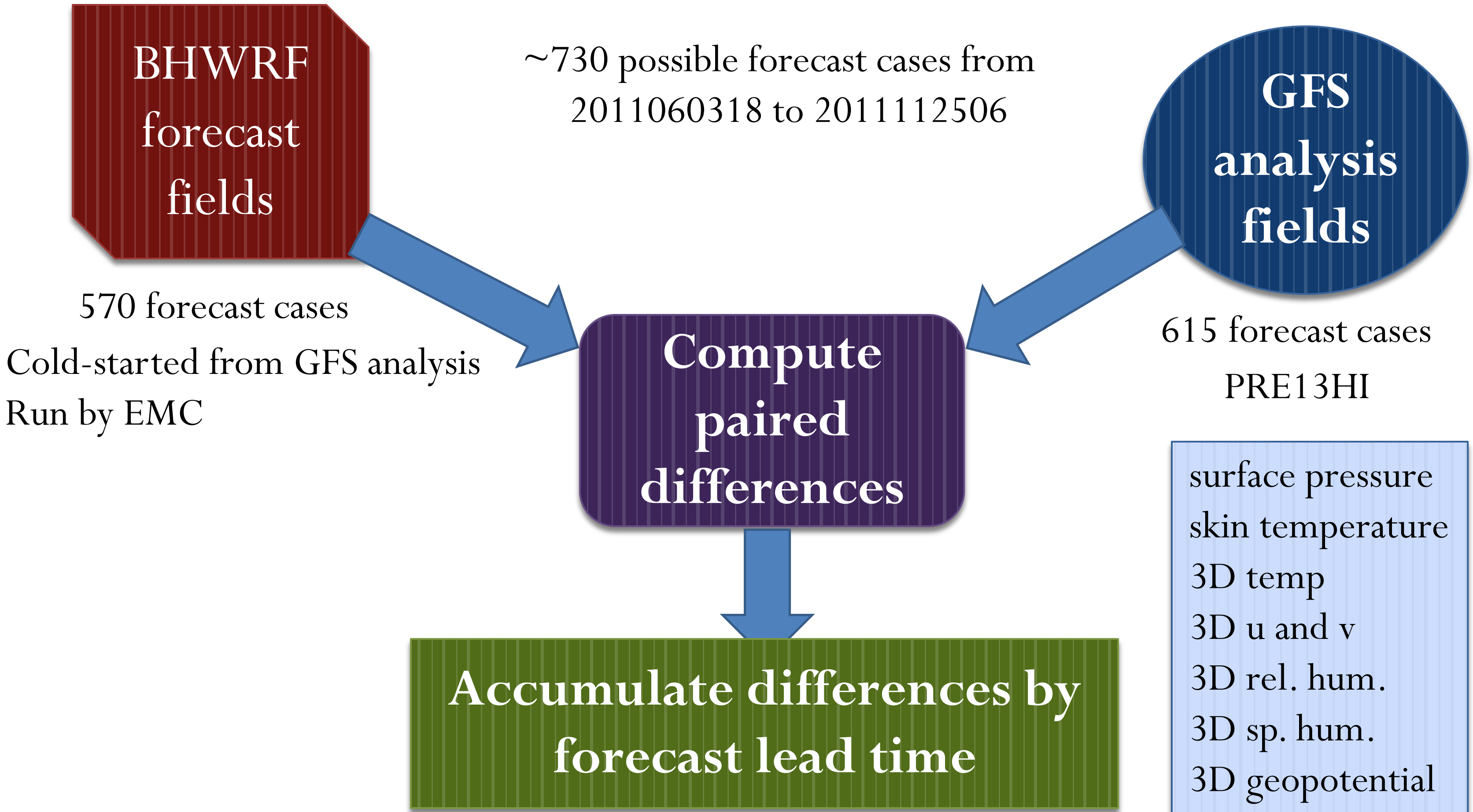
Background

- **Motivation**
  - EMC is preparing to implement basinscale HWRF in '14/15
  - Extensive collective work in data assimilation, moving nests, trans-Atlantic POM
  - Need to identify large scale errors – Vx of HWRF 3D fields never done before
- **DTC diagnostic study**
  - Evaluated cold-started basinscale HWRF large scale fields
  - Identified issues that deserve further investigation (hypotheses)
  - Created benchmark

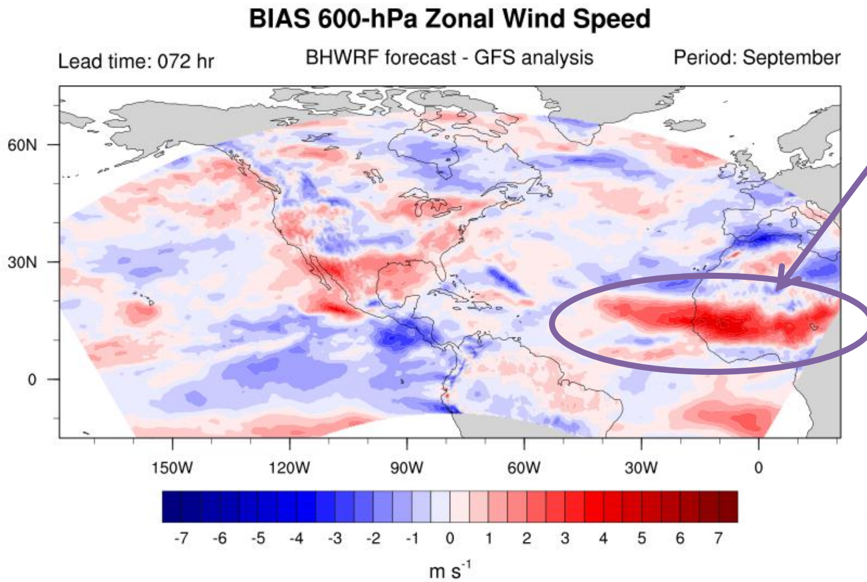
Example of basinscale domain



Methodology



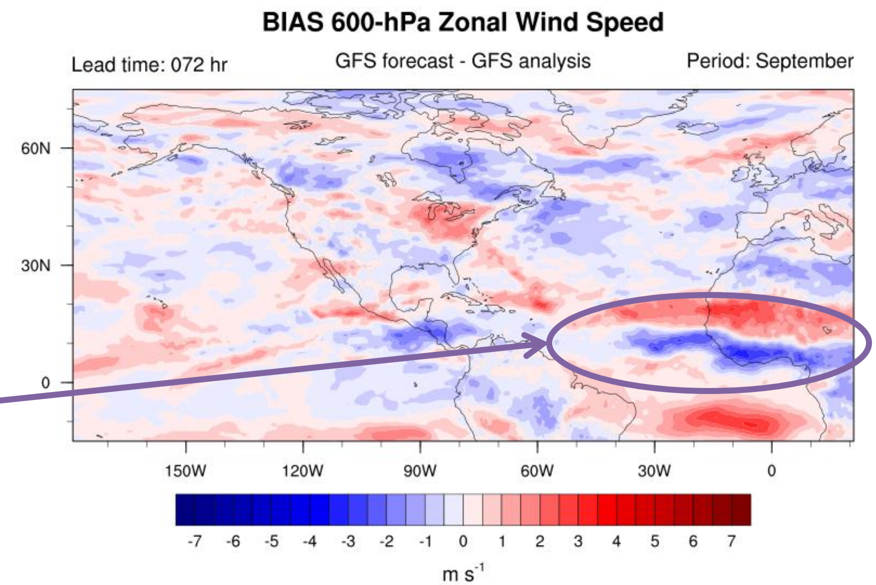
Highlight: 600-hPa zonal wind speed



Basinscale bias

September 2011 – 72-h forecast

African jet too weak in HWRF



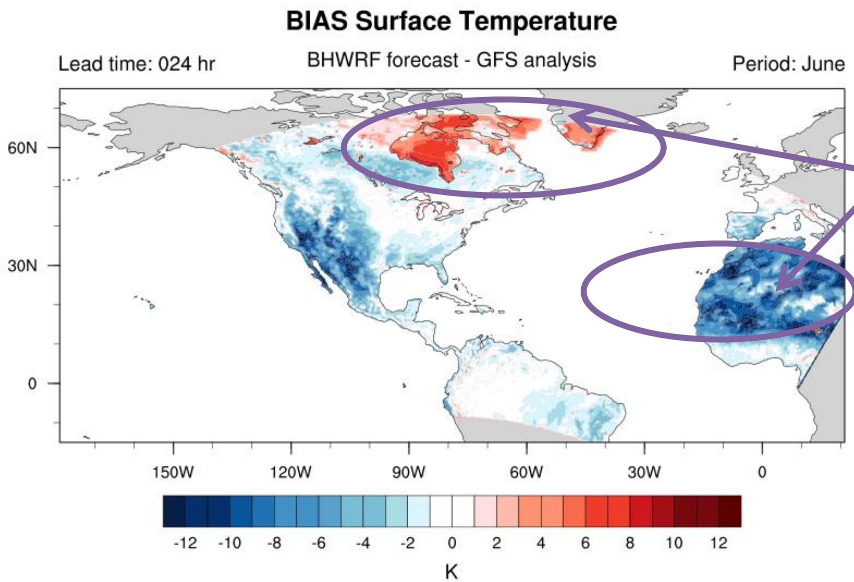
GFS Bias

September 2011 – 72-h forecast

In GFS jet displaced to south

Developmental Testbed Center

Highlight: surface temperature

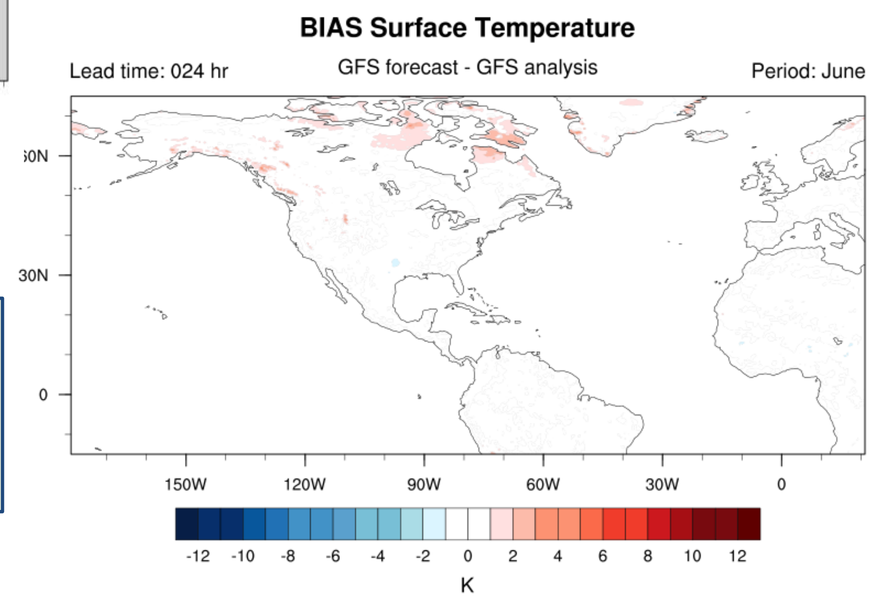


Basinscale bias

June 2011 – 24-h forecast

HWRF cold over dry continental areas

Suggests issue with inland ice



GFS Bias

June 2011 – 24-h forecast

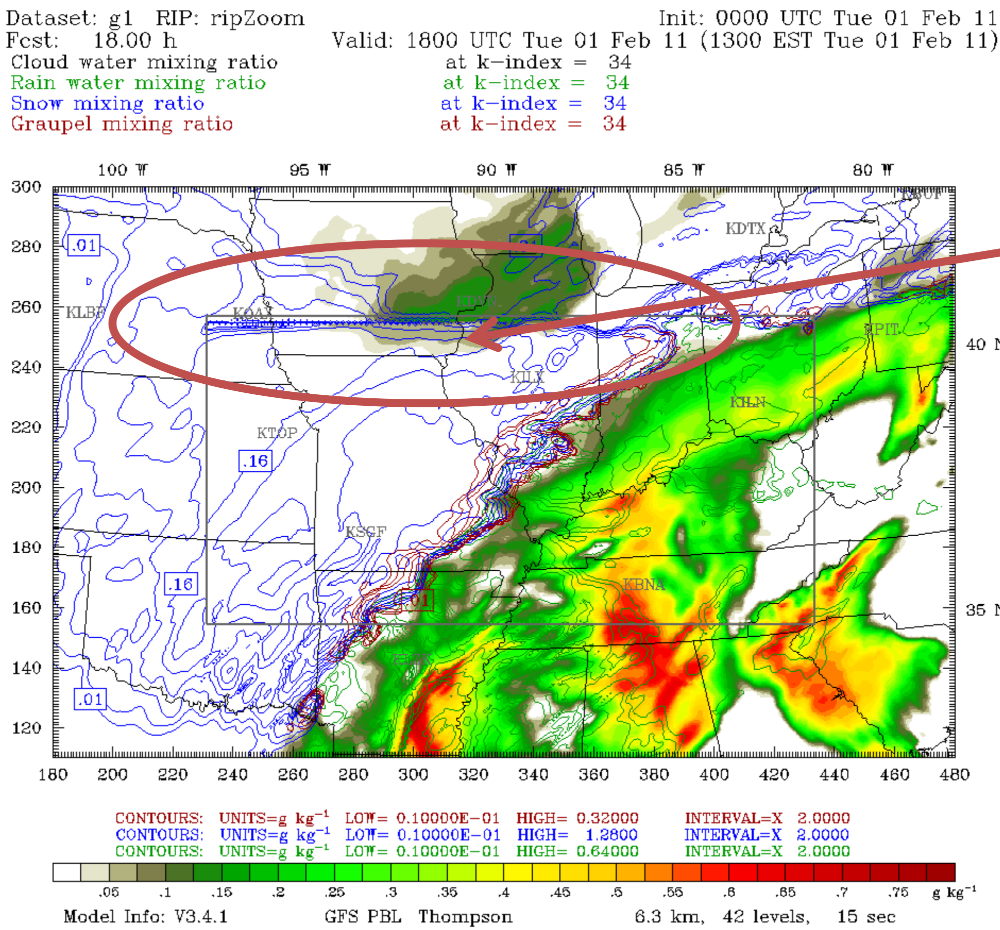
No significant biases

Thompson microphysics

DTC-EMC collaboration in MP

- **Interoperability**
  - EMC (S. Trahan) has created the basic interoperability
    - Ability to advect various microphysics mixing ratios and number concentrations (Ferrier only advects one species)
    - New nest-parent interpolation routines which communicate all microphysics variables (for Ferrier or other microphysics)
  - DTC improving MP-radiation interface
- **Testing by DTC**
  - Irene and Earl, with stationary and moving nests
  - Winter storm with single domain and stationary nest
- **Debugging**
    - Tests, diagnostics, code analyses uncovered bugs in nest-parent interpolation
    - EMC corrected; work in progress

HWRF w/Thompson MP (winter storm)



Most recent problem solved: snow coming from grid1 into grid2 has a sharp discontinuity (also cloud ice number concentration).

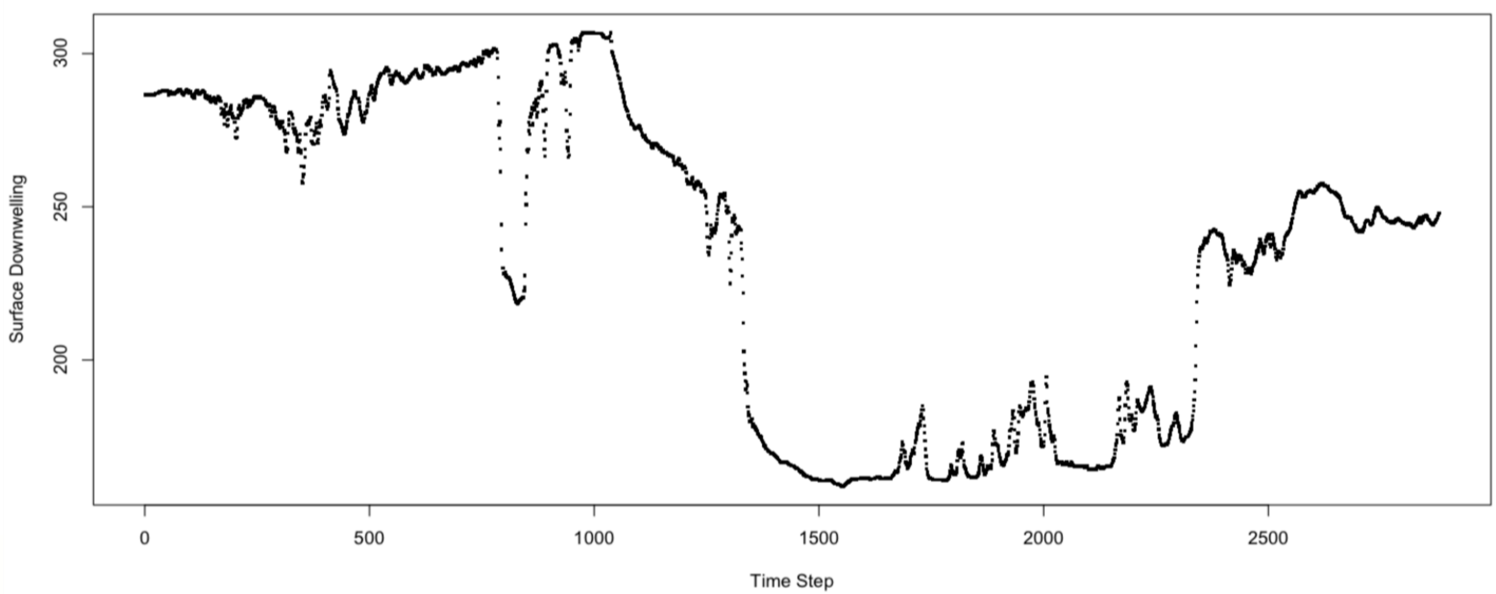
Caused by an array dimensioned incorrectly

Radiation code issues: DTC work

- The sum of ice and snow mass is passed from MP to radiation
- Their radius is assumed to be small at cold temperatures
- Effectively, snow is counted as small particles, with massive (and incorrect) impact on shortwave radiation reflection
- Solution: compute effective radii of cloud ice, snow, cloud droplets in manner consistent with microphysics scheme – for Thompson, Ferrier etc.
- Implemented in WRF-ARW in RRTMG (RRTMG being tested by EMC for 2013 HWRF)
- Will transfer to HWRF ***and*** NMM-B

Leveraging SURFRAD in MET

- SURFRAD ingest in METv4.1
  - Useful for radiation scheme evaluations
  - Land surface model verification
  - Solar forecast evaluation for DOE project



BACKGROUND